Acta Cryst. (1998). D54, 1023–1025

Crystallization of succinylated concanavalin A bound to a synthetic bivalent ligand and preliminary structural analysis

DAVINA N. MOOTHOO,[a] STEPHEN A. MCMAHON,[a] SARAH M. DIMICK,[b] ERIC J. TOONE[b] AND JAMES H. NAISMITH[a]* at [a]*Centre for Biomolecular Science, Purdie Building, The University, St Andrews KY16 9ST, Scotland, and* [b]*Departments of Biochemistry and Chemistry, Duke University, North Carolina, NC 27708-0346, USA. E-mail: naismith@st-and.ac.uk*

(*Received 2 February 1998; accepted 9 March 1998*)

Abstract

Crystals have been obtained of succinylated concanavalin A complexed to a novel bidentate synthetic ligand. The crystals are the first example of a lectin with a synthetic multivalent ligand and the first report of crystallization of succinylated concanavalin A. The crystals were obtained by sitting-drop vapour diffusion equilibrating with a solution of 20% polyethylene glycol, pH 5, 293.5 K. Crystals are orthorhombic, belonging to space group $C222_1$ with unit-cell dimensions of $a = 99.1$, $b = 127.4$, $c = 118.9$ Å. The asymmetric unit contains a dimer, with over 65% of the volume occupied by water. The ligand cross links concanavalin A monomers. Succinylated concanavalin A is known to be a dimer in solution, yet it is found as the typical concanavalin A tetramer in the crystal. The contacts holding together the tetramer appear extensive and suggest that a fine balance between dimer and tetramers exists. Data to 2.65 Å have been collected and the structure determined by the molecular replacement method.

1. Introduction

Lectins comprise a varied family of sugar-binding proteins; they are found in all types of organisms. Plant lectins in particular have been of intense interest, because they exhibit exquisite specificity for oligosacccharides and unlike their mammalian counterparts are much more tractable to characterization by a broad range of biophysical techniques. Thus, these proteins, such as the lectin from *Canavalia ensiformis*, concanavalin A (con A), have served as models for more complex species (Rini, 1995). In contrast to the exquisite selectivity at the oligosaccharide level, the proteins are relatively promiscuous at the monosaccharide level, falling into two broad classes; mannose specific and galactose specific. The affinity of the lectin for monosaccharides is correspondingly lower, of the order $K_a = 1 \times 10^3\ M^{-1}$ compared with $K_a = 1 \times 10^6\ M^{-1}$ for oligosaccharides (Chervenak & Toone, 1995; Mandal *et al.*, 1994; Toone, 1994). The advent of calorimetric data on lectin (especially con A) carbohydrate complexes and the realisation that *ab initio* modelling methods were failing to accurately model protein–carbohydrate interactions has reinvigorated the structural study of lectin–carbohydrate complexes. A number of oligosaccharide–protein complexes have been determined (Rini, 1995), most recently the structures of con A with its cognate trisaccharide and pentasaccharide (Naismith & Field, 1996; Moothoo & Naismith, 1998).

Protein–saccharide interactions are appealing therapeutic targets in many diseases, particularly those involving infection and inappropriate immune response (Dwek, 1996). Therapeutics designed for this task would principally be aimed at disrupting or interfering with molecular recognition processes rather than being cytotoxic. Oligosaccharides are poor therapeutics, however, as they are too polar for satisfactory uptake. Monosaccharides, however, bind too weakly to be of value. One possible solution to this conundrum has been the use of polyvalent ligands. This is based on the observation that in nature many carbohydrate proteins appear to function as oligomers. Chemical synthesis has produced families of such multivalent ligands, which on the basis of agglutination assays appear to have dramatic results (Kanai *et al.*, 1997; Sigal *et al.*, 1996). Several structures are now known of lectins complexed to naturally occurring carbohydrates that cross link the protein in the crystalline phase (Bourne *et al.*, 1994; Dessen *et al.*, 1995; Wright, 1992; Wright & Hester, 1996). Important questions remain as to the thermodynamic basis of these polyvalent interactions (Roy, 1996; Toone, 1994). As part of a program to combine calorimetery, crystallography and chemical synthesis we report the crystallization and structure determination of the first synthetic multivalent ligand–lectin complex.

2. Crystallization and X-ray data collection

Succinylated con A was purchased from Sigma (Poole, UK). The synthesis of the bidentate ligand [1,3-di-(*N*-propyloxy-α-D-mannopyranosyl)-carbomyl 5-methylazido-benzene, Fig. 1] will be described elsewhere. Crystallization trials were performed by the means of the sitting-drop vapour-diffusion method (Ducruix & Giegé, 1992). A solution of succinylated con A (1.2 m*M*)/ligand (18 m*M*) was prepared in 20 m*M* Tris pH 7, 100 m*M* NaCl, 1 m*M* $CaCl_2$ and 1 m*M* $MnCl_2$ and equilibrated against 20% polyethylene glycol (M_r = 6000), 100 m*M* citric acid, pH 5 in sitting-drop trays (Charles Supper, USA) at 293.5 K. Small crystals of dimensions 0.1 × 0.2 × 0.2 mm were initially obtained. Optimization of the ligand concentration (5 m*M*) yielded block-shaped crystals of up to 1.0 × 0.8 × 0.6 mm. Crystal growth was complete in 14 d. 12 crystals were examined, mounted in a thin-walled glass capil-

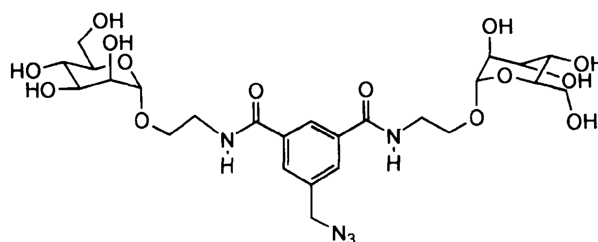

Fig. 1. The bivalent ligand [1,3-di-(*N*-propyloxy-α-D-mannopyranosyl)-carbomyl 5-methylazido-benzene] which cross links con A.

© 1998 International Union of Crystallography
Printed in Great Britain – all rights reserved

ISSN 0907-4449 © 1998

lary, by exposure to X-rays. The resolution limit varied from 2.65 to 3.5 Å. All diffraction data were collected at 293.5 K from a single crystal (0.6 × 0.5 × 0.3 mm) which diffracted to the highest resolution. Larger crystals did not give higher resolution data. All data were recorded on the Nonius/MacScience DIP2000 dual image plate. X-rays were generated using a Enraf–Nonius FR951 rotating-anode generator and focused using the MacScience mirror system, through a 0.5 mm collimator. Data were collected as 92 non-overlapping 25 min 1° oscillations with a crystal-to-detector distance of 140 mm. The resolution limit of the data was 2.65 Å and no significant crystal decay was observed during data collection. The programs *DENZO* and *SCALEPACK* (Otwinowski, 1993) were used to process the data. The crystal was indexed in a centred orthorhombic space group with unit-cell dimensions $a = 99.1$, $b = 127.4$, $c = 118.9$ Å. Analysis of diffraction data identified systematic absences consistent with space group $C222_1$. A dimer of molecular mass 49 kDa, gives rise to a V_m (Matthews, 1968) of 3.8 $Å^3\,Da^{-1}$ and indicates a solvent content of over 65%. The data are 96% complete from 26 to 2.65 Å, with an R_{merge} 7.2% with a redundancy of 2.2 and a total of 84% of possible data are greater than 1σ. For the high-resolution shell (2.75–2.65 Å), the corresponding values are 98% complete, R_{merge} 20.2%, redundancy 2.2 and 83% greater than 1σ.

3. Structure solution

The structure was solved by molecular replacement, using *AMoRe* (Navaza, 1994) as implemented in the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). The conventional con A dimer (monomers *A* and *B*) from the trimannoside complex of con A (1CVN) stripped of metals, waters and sugars was used as the search model. The rotation function found a single solution with a correlation coefficient of 0.28 and the translation function produced a final solution with a correlation coefficient of 0.84. A translation search in *C*222 gave no solution. In order to visualize the interaction of the ligand with each protein monomer, the asymmetric unit was redefined such that one monomer of the con A dimer (monomer *A*) was linked *via* the ligand to another monomer (monomer *C*). Monomer *C* was generated by applying the symmetry transformation ($\frac{1}{2} - x$, $\frac{1}{2} - y$, $z + \frac{1}{2}$) to monomer *B* (Fig. 2). Other crystallographic operators generate the 'typical' con A tetramer. The initial free *R* factor was 38% to 2.65 Å. Rigid-body refinement of the monomers gave a free *R* factor of 30%. Clear density is visible for the metal ions and ligand. The ligand is currently being built into this electron density and a stereochemical dictionary constructed. Once the ligand has been built into density, refinement will recommence. A fully refined structure will be reported elsewhere.

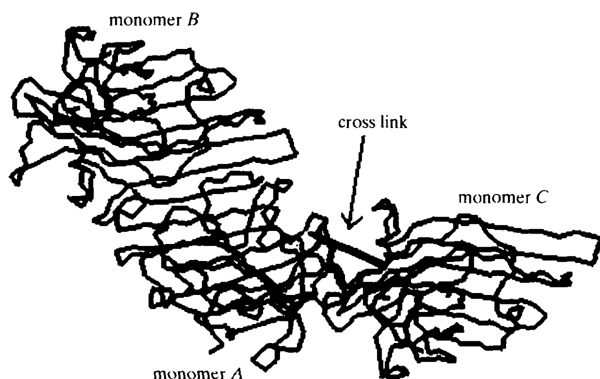

Fig. 2. The conventional dimer of con A is formed by monomer *A* and monomer *B*, the ligand cross links monomer *A* to monomer *C*. Monomer *C* is related by the crystallographic transformation ($\frac{1}{2} - x$, $\frac{1}{2} - y$, $z + \frac{1}{2}$) to monomer *B*. The cross-linking ligand is shown a straight line between the two sugar-binding sites.

That succinylated con A is found as a tetramer in this crystal is puzzling; the electron density shows evidence of succinylation on two lysine residues. Dimerization of the dimers occludes 4300 $Å^2$ of surface area which compares to 5200 $Å^2$ for native con A (Deacon *et al.*, 1997; Reeke *et al.*, 1975) and 4800 $Å^2$ for the mannose con A complex (Naismith *et al.*, 1994). There are 120 protein–protein contacts less than 4.0 Å between the two dimers in this structure, compared with 260 for the native and 160 for the mannose complex. Although this tetramer is clearly less tightly packed than the other two, by a standard set of standard crystallographers' criteria (buried surface area and contacts) one would identify the protein to be a functional tetramer. However, solution studies show unambiguously that succinylated con A is a dimer (Gunther *et al.*, 1973). It appears that the energy barrier between dimers and tetramers for succinylated con A is small enough that crystal-packing forces can drive formation of the tetramer. We feel this result poses an interesting test for those who wish to devise automatic methods of identifying protein oligomerization state from static crystal structures.

JHN acknowledges BBSRC support (B08307) and we thank John Helliwell for encouragement.

References

Bourne, Y., Bolgiano, B., Liao, D. L., Strecker, G., Cantau, P., Herzberg, O., Feizi, T. & Cambillau, C. (1994). *Nature Struct. Biol.* **1**, 863–870.

Chervenak, M. C. & Toone, E. J. (1995). *Biochemistry*, **34**, 5685–5695.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Deacon, A., Gleichmann, T., Kalb (Gilboa), A. J., Price, H., Rafferty, J., Bradbrook, G., Yariv, J. & Helliwell, J. R. (1997). *J. Chem. Soc. Faraday Trans.* **53**, 4305–4312.

Dessen, A., Gupta, D., Sabesan, S., Brewer, C. F. & Sacchettini, J. C. (1995). *Biochemistry*, **34**, 4933–4942.

Ducruix, A. & Giegé, R. (1992). *Crystallization of Nucleic Acids and Proteins*. Oxford: IRL Press.

Dwek, R. A. (1996). *Chem. Rev.* **96**, 683–720.

Gunther, G. R., Wang, J. L., Yahara, I., Cunningham, B. A. & Edelman, G. (1973). *Proc. Natl Acad. Sci. USA*, **70**, 1012–1016.

Kanai, M., Mortell, K. H. & Kiessling, L. L. (1997). *J. Am. Chem. Soc.* **119**, 9931–9932.

Mandal, D. K., Kishore, N. & Brewer, C. F. (1994). *Biochemistry*, **33**, 1149–1156.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Moothoo, D. N. & Naismith, J. H. (1998). *Glycobiology*, **8**, 173–181.

Naismith, J. H., Emmerich, C., Habash, J., Harrop, S. J., Helliwell, J. R., Hunter, W. N., Raftery, J., Kalb, A. J. & Yariv, J. (1994). *Acta Cryst.* D**50**, 847–858.

Naismith, J. H. & Field, R. A. (1996). *J. Biol. Chem.* **271**, 972–976.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Otwinowski, Z. (1993). In *Proceedings of the CCP4 study weekend: Data collection and processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.

Reeke, G. N. J., Becker, J. W. & Edelman, G. M. (1975). *J. Biol. Chem.* **250**, 1525–1547.
Rini, J. M. (1995). *Annu. Rev. Biophys. Biomol. Struct.* **24**, 551–577.
Roy, R. (1996). *Curr. Opin. Struct. Biol.* **6**, 692–702.
Sigal, G. B., Mammen, M., Dahmann, G. & Whitesides, G. M. (1996). *J. Am. Chem. Soc.* **118**, 3789–3800.
Toone, E. J. (1994). *Curr. Opin. Struct. Biol.* **4**, 719–728.
Wright, C. S. (1992). *J. Biol. Chem.* **267**, 14345–14352.
Wright, C. S. & Hester, G. (1996). *Structure*, **4**, 1339–1352.